



# Kent Academic Repository

**Watt, John, Sinnott, Richard O., Inman, George and Chadwick, David (2011)**  
***Federated Authentication and Authorisation in the Social Science Domain.***  
**In: 2011 Sixth International Conference on Availability, Reliability and Security.**  
**IEEE, pp. 541-548. ISBN 978-1-4577-0979-1.**

## Downloaded from

<https://kar.kent.ac.uk/31978/> The University of Kent's Academic Repository KAR

## The version of record is available from

<https://doi.org/10.1109/ARES.2011.83>

## This document version

UNSPECIFIED

## DOI for this version

## Licence for this version

UNSPECIFIED

## Additional information

## Versions of research works

### Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site.  
Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

## Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

# Kent Academic Repository

## Full text document (pdf)

### Citation for published version

Watt, John P. and Sinnott, Richard O. and Inman, George and Chadwick, David W. (2011) Federated Authentication and Authorisation in the Social Science Domain. In: 2011 Sixth International Conference on Availability, Reliability and Security (ARES 2011), 22-26 Aug 2011, Vienna.

### DOI

<https://doi.org/10.1109/ARES.2011.83>

### Link to record in KAR

<http://kar.kent.ac.uk/31978/>

### Document Version

UNSPECIFIED

#### Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

#### Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

#### Enquiries

For any further enquiries regarding the licence status of this document, please contact:

**[researchsupport@kent.ac.uk](mailto:researchsupport@kent.ac.uk)**

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

## Federated Authentication and Authorisation in the Social Science Domain

John Watt  
National e-Science Centre  
University of Glasgow  
Glasgow G12 8QQ, UK  
Email: j.watt@nesc.gla.ac.uk

Richard O. Sinnott  
Melbourne eResearch Group  
University of Melbourne  
Melbourne, Australia  
Email: rsinnott@unimelb.edu.au

George Inman, David Chadwick  
Information Systems Security Research Group  
University of Kent  
Canterbury, UK  
Email: G.Inman@kent.ac.uk

**Abstract**—The use of Shibboleth as a mechanism for implementing federated authentication is commonplace in many countries. The ability of Shibboleth to transmit extra information about a user, including licenses, roles and other attributes, is not exploited for many reasons, mainly because institutional Identity Providers (IdPs) are not maintainable sources of fine-grained authorisation information. The JISC-funded Shintau project has produced an extension to the Shibboleth profile which allows a user to link information from more than one IdP together utilising a custom Linking Service (LS). This paper describes both the application and independent evaluation of this software by the National e-Science Centre (NeSC) at the University of Glasgow within the context of the ESRC-funded Data Management through e-Social Science (DAMES) project.

**Keywords**—Shibboleth, Shintau, authorisation, attribute aggregation, SAML

### I. INTRODUCTION

The e-Social Science domain is providing numerous challenges in secure data access and management. Some of the most important resources for studying the lives of citizens require stringent technical procedures to be followed in order to gain access, often with some sort of legal obligation that needs to be enforced. This is especially true with regard to the historic primary source of a nation's demographic - a full, national census [1]. In the UK these are carried out every ten years, with the next (and allegedly the last) scheduled for 2011. The results of the census may be made public with multiple granularities of access and usage, from key statistics that paint a picture of general trends across the country that are freely available for download, down to individual-level microdata which may be regarded as "disclosing" and thus should only be made available under strict terms and conditions. The typical end-user experience for acquiring this data starts with a visit to the data centre's web site. Here it will be necessary to register a username, normally a valid email address, and password as an authentication token. Sometimes the only check on this token is to look up the fact that the email is valid, or is an email from an academic institution. Next the user will be required to select which data sets are of interest to them out of a list of available downloads. The next step is for the data centre to perform an authorisation, which takes the form of an

End User License Agreement (EULA) type statement which the user is required to accept, specifying the conditions upon which the data is offered and any obligations, legal or otherwise, which the user is required to undertake. For more disclosing data sets, this download is currently limited, with many organisations insisting on extra restrictions on access and usage. This can for example through demanding the user physically attend a secure location where this data can be accessed and used (subject to numerous previous checks justifying why this access is needed). They are often monitored when such data access and usage is undertaken. This, whilst offering a secure access and usage model places obvious constraints on the end user community and thus minimizes the research possibilities that these data sets provide.

Figure 1 shows an example of the special conditions which are available at one such data provider - the Casweb data archive [2]. These restrictions state, for example, how securely the user will have to store the downloaded data, and also whether results extracted from these data sets may be externally published. One of the key challenges with these procedures is how to make the agreements or licences that a user has signed up to available for an external automated process such as a web portal. This is especially so when different data sets from different data providers each with their own security and licensing agreements, need to be supported and adhered to. This is a common occurrence for researchers wishing to access individual-level social, clinical and other more sensitive data resources for example. Thus whilst the model described above is essentially interactive, requiring a user visit a particular web site, sign up to the terms and conditions, then download and store the data themselves, a far better solution is to support user-oriented access to and usage of a multitude of security-oriented data resources from a range of data providers, where each provider is still autonomous and able to define and enforce the local terms and conditions on access and usage of their data sets. This is a key demand from data providers in the Data Management through e-Social Science (DAMES) [3] project.

User management for large-scale projects or resources carries large administrative burdens on collaborating centres.

The establishment of a user's identity, the maintenance of that identity and the allocation and revocation of privileges and licenses comprise a heavy workload that requires dedicated persons to be employed at institutions. If centres insist on maintaining credentials for any external users who access their systems, there will come a point where the administrative workload will become too great. In addition, the authenticity of these users is hard to establish if they are based at remote institutions. In the past ten years, there has been a paradigm shift towards the concept of *federated* identity and access management infrastructures. This allows a group of establishments such as higher education institutions to agree to form a federation of trusted sites who place trust in the identity assertions of the individual collaborating sites. From an end user perspective, this allows a user wishing to access a service/resource available through the federation to login using only their local credentials. Once authenticated, these credentials can subsequently be used to access other resources available through the federation without further authentication challenge/responses - known as Single Sign-On (SSO). Numerous gains are made through collaboration based on such federations: institutions will only have to manage their local users as they have always done, plus end users will only have to manage one set of credentials to access a whole range of resources which would otherwise need their own usernames, passwords etc. In the United Kingdom, the UK Access Management Federation [4] is responsible for registering entities and maintaining metadata. These infrastructures work well when all the home institution needs to do is release an authentication statement about a user. Releasing extra information, in particular what privileges or licenses they hold on other systems is a more complicated scenario, made difficult by the fact that licenses are often held by external authorities (generally the resource host) and the management of user authorisation which may change on a day-to-day basis is not something that home

institutions is rightly able to undertake.

The ESRC-funded DAMES project is a three-year project looking at social science data management activities relating to occupation, education, ethnicity and clinical/e-Health and wider social care data sets. Data management challenges faced in the social sciences are numerous: investigation of trends in data over decades (longitudinal studies) where different coding systems and classifications and categorisations have been defined and used by the community, e.g. regional boundaries of local authorities change over time and understanding this when dealing with changing population dynamics or the change/impact upon local, regional and national health policy, is essential to guide policy and aid understanding of issues impacting upon society more generally. Many of these data challenges implicitly require interfacing and access to data from other disciplines, e.g. the clinical domain, the geospatial domain, economic domain etc. As noted access to many of the key data sets in this context is especially fraught, where restrictions and obligations on the user is mandatory. Solutions that help to simplify this overall process for end users and capturing the autonomy and access control demanded by the data providers and associated stakeholders themselves are thus highly sought after.

This paper describes the utilisation of the recently released ShinTau middleware [5] that enables a user to create a linked set of external licenses, permissions and roles that will allow them to access multiple secure data sets available through the DAMES web portal adhering to a strict security policy. This infrastructure aims to augment federated authentication now widely adopted across the community through the UK Access Management Federation for example, with federated authorization whereby users are able to manage multiple credentials from multiple authorities in a seamless manner. A key motivation behind this is to minimize the effort required by both the end users and the data providers themselves.

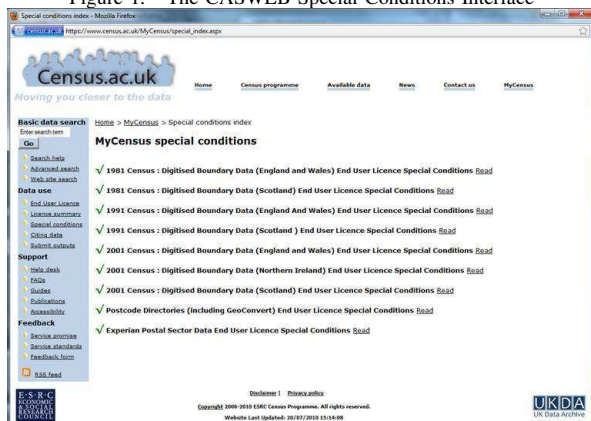
## II. TECHNOLOGIES

The ShinTau infrastructure relies on several widely available user and access management technologies. These will be described in the following section, with the final subsection describing the ShinTau software itself.

### A. Shibboleth

The Shibboleth [6] implementation of the Security Assertion Markup Language (SAML) has become the *de facto* standard in implementing federated access control to web resources, particularly in the higher/further education domain. In exploiting this software, national-level resource providers may delegate their user authentication (and trust) requirements to their national federation. This federation is responsible for registering and maintaining entities who wish to collaborate in this trust network, ensuring that Service Providers (SPs) have access to fully accountable

Figure 1. The CASWEB Special Conditions Interface



Identity Providers (IdPs). This infrastructure works well when resources require merely a confirmation of a user's status within an organisation, or a statement of authenticity that ensures the person accessing the service is who they say they are.

The SAML specification allows for extra information to be asserted about an individual, so that further decisions on their privilege on a given resource may be made. The eduPerson specification [7] defines a set of standard user attributes which federations may adopt to ensure a basic minimum of information is released about a user for proper accountability and auditing. In the UK, institutional IdPs have been found to reliably release only two of the eduPerson attributes, namely eduPersonAffiliation and eduPersonTargetedId. The former attribute is a scoped statement of a user's status within the institution they have credentials for, for example, a student at the University of Glasgow would find this attribute populated by the value student@ gla.ac.uk. Scoped, in this sense, means the location-specific information appended to the attribute - in this case ' gla.ac.uk'. Any SP receiving this attribute can be assured (to the extent of the trust model underlying the federation itself) that the person logging in using the credentials that generated the assertion is a student of the University of Glasgow. The latter attribute mentioned above, eduPersonTargetedId, is a non-disclosing hash that has a unique and stateful value for a particular user trying to access a particular SP logging in from a particular IdP. This attribute may be used to identify a returning user to a resource without disclosing information such as first name, date of birth, etc. In this way, anonymous access to web resources is possible, because whilst the SAML assertion confirms the identity of the user, the assertion contains no disclosing information.

The eduPerson specification defines several other attributes for transmission, but these are not well supported across all IdPs in a federation due to concerns about privacy. One of interest is eduPersonEntitlement, which is an attribute which may take values which can be used to authorise further actions on a resource. For example, a portal may request that an entitlement value which represents 'portal administrator' should be asserted before access to specific content is granted. The main difference between this attribute and the two eduPerson attributes described previously is that eduPersonEntitlement is likely to contain values that may change fairly frequently, as resources with different requirements come and go or alter their demands. The eduPersonAffiliation and eduPersonTargetedId attributes are, by contrast, essentially static based on whatever database is providing the user information. So administration of these two common attributes is easy and supported by nearly all IdPs, however support for eduPersonEntitlement is generally not given because of the administration and accounting costs. In order to release this attribute, sites would have to manage the changeable, day-to-day authorisation tokens required by

users on systems that they themselves have no authority to issue. Thus the University of Glasgow should not issue authorization tokens granting access to the UK Census for example, unless specifically delegated the task by the UK Census data providers. Software to support such delegation of authority was described in [8] and applied in [9].

## B. PERMIS

PERMIS [10] is a generic authorisation infrastructure which issues roles and privileges to users using X.509 [11] Attribute Certificates and subsequently uses them to enforce authorisation policy. PERMIS provides a plug-in authorisation enforcement point (PEP), and tools with which to issue ACs and write local security policies. A security policy is a document (typically written in XML) that details the precise access control requirements of a resource. In PERMIS, policy definition and enforcement use an XML triple comprising a Role, Action & Target. For many purposes, the Role is contained within an X.509 Attribute Certificate (AC) [12] which the user provides directly or through extraction of the Distinguished Name (DN) and subsequent LDAP lookup. The Target represents the URI of the Grid/Web service which the user is attempting to access, and the Action is the individual method that the user is attempting to invoke on this target. The XML policy dictates which combinations of Action and Target are permitted based on the Role that the user presents in their AC. Additional rules about which PKI keypairs are recognised by the PERMIS PDP, which certificate DNs are permitted and their validity time can be expressed in the XML policy. All objects in PERMIS are digitally signed, ensuring the information hasn't been tampered with, and allowing the issuer of the certificate/policy to be confirmed. PERMIS also supports associated tools to help in this process including an Attribute Certificate Manager and Policy Editor.

The PERMIS module used for the ShinTau demonstration is the PERMIS Shibboleth Apache Authorisation Module (SAAM) [13], which is a plug-in for Apache Web Server that allows locations within the Apache configuration to be protected by PERMIS. A modified version of the PERMIS decision engine was used as the Policy Information Point responsible for collecting attributes from the remote IdPs.

## C. ShinTau

The Information Systems Security Research Group at the University of Kent have developed the ShinTau [14] infrastructure for enriching the attribute sets that a vanilla Shibboleth installation provides, addressing the concerns over privileges or licenses required for external resource access as discussed in the previous section. The Shintau architecture was informed by a requirements gathering exercise during 2007 [15]. The implementation of ShinTau consists of a standard Shibboleth SP protecting web content hosted in Apache Web Server, with an extra directive to

allow authorisation on the protected area to be performed by a policy-enforcing PERMIS application. Note here that the SP is the only unmodified Shibboleth entity in the infrastructure. In order to achieve user-instigated attribute linkage, a new entity within the SAML profile is required called a Linking Service (LS). This is a PHP-based web application that allows users to log in to (ShinTau-enabled) IdPs and create mappings between SPs and multiple IdPs. The Linking Service contains a novel metadata entry that behaves like a hybrid IdP and SP, unfortunately the use of the extra extensions (Liberty Alliance End Point Reference) means that the UK Access Management Federation refuse to add entities like this (and the modified IdPs) to their federation metadata. Because of this, the testshib.org test federation was used for all implementation work.

The Shintau infrastructure calls for the user to first create the IdP mappings in the LS. This requires the user to have some knowledge of which attributes are required for a particular service. In some cases the SP will inform the user of the access requirements, and in some this may be less visible. The point being that the *user* is the person in control over the composition of the attribute set that is finally presented to the SP. It is important to note that at this stage the user cannot dictate precisely which attributes are presented by a single IdP, this is beyond the scope of influence of the user and is set in the IdP's Attribute Release Policy which is managed by the IdP's administrator. Rather, the user can manage the union of multiple IdPs to create a merged set of capabilities, which may include permissions beyond those which the SP requires. For example, it would be expected that a portal making use of census data would not require access to *every* piece of data that the user had signed the agreements to. In the case of an institutional IdP this problem is unlikely to be a factor, for the administrative reasons outlined earlier. However, a data provider making its user licenses available to SPs through a custom IdP will probably populate an attribute assertion with every license the user holds as there exists no current mechanism to select these attributes other than for the SP to filter the incoming statement accordingly.

IdPs are accessed through the Shintau Linking Service's "Link Account" function. This function redirects the user to the Shibboleth Where Are You From (WAYF) service where the user is prompted to select the IdP of interest from a list of the federation members. Upon selection, the user is redirected to the IdP to login using the credentials required at that site. Once the user has successfully authenticated at the IdP, they are returned to the Linking Service, where the persistent details of that login session are stored in a database and displayed to the user in a table. Each subsequent authentication to a federation IdP will result in the persistent details being fetched from that IdP and added to the linkage table. Once the user has logged in to all the IdPs they need to extract information from, the next step is to associate all

or a subset of these IdPs with a particular federation SP. The Linking Service "Release Policy" configuration function is used for this purpose. This function allows a user to select all or a subset of the IdPs they have just registered and map them as being available to an individual federation SP (which will typically require authorisation information available from multiple sources of authority). This can be done for as many SPs as the user requires.

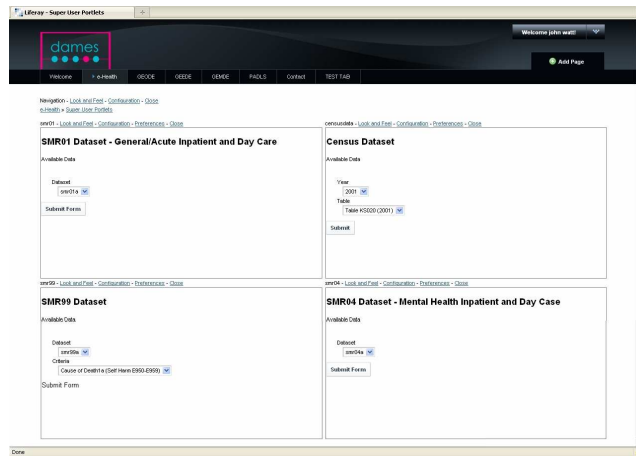
Once the link between multiple IdPs and SP has been made, secure information from multiple sources of authority (IdP) is available for a PERMIS service running on the SP to access the IdPs attribute assertions through the Liberty Alliance Discovery Service. This action is triggered when visiting the SP, where the user is immediately sent to the WAYF to select their primary IdP (usually an institutional IdP with high user accountability). The user provides their credentials as normal, but before submitting them, they are offered an extra check box on the IdP login page to "Aggregate Attributes". This tells the user that the final attribute assertions will not all come directly from this IdP, but some will come from the Linking Service which fetches the authorization credentials available to the SP being accessed and checks the user mappings to see which IdPs have been marked as sources of user information for this particular SP. The Linking Service then coordinates the attribute retrieval from each IdP for the PERMIS SAAM service to subsequently use to enforce access and authorization policy, e.g. to check the validity of the attributes and fulfillment of all local access and usage policies associated with the SP. If the correct merged attribute set has arrived at the SP, the PERMIS PDP grants access to the SP. If the policy has not been satisfied by the attribute aggregate, the user is denied access. In this manner, multiple user-selected attribute authorities can be used to make single access control decisions to providers requiring attributes from a variety of sources of authority. This model is essential in security-focused domains where data providers demand finer-grained access control.

### III. DAMES USE CASE

The ESRC-funded DAMES project has been ongoing since January 2008. The project has developed a portfolio of specialist data environments focused on specific social science themes:

- Grid-enabled Occupational Data Environment (GEODE) - which has developed an environment for research into the data challenges associated with occupational data resources [16];
- Grid-enabled Minority Data Environment (GEMDE) - for data challenges associated with ethnicity and minority related data resources - [17];
- Grid-enabled Education Data Environment (GEEDE) - for data challenges associated with education and qualification related data resources [18];

Figure 2. The DAMES portal e-Health Portlets for Depression, Self Harm and Suicide Related Research



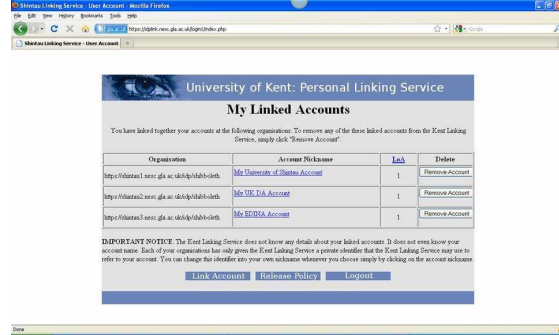
- e-Health specialist data environment - which has established an environment for a range of e-Health related research topics with specific focus on inter-disciplinary research challenges.

The DAMES e-Health related environment has acute demands on security, especially since it is dealing with data from organisations such as the National Health Service, the Census and license protected geospatial data sets. The e-Health environment has focused in particular on supporting research into depression, self-harm and suicide.

As shown in Figure 2, several portlets are available within the DAMES portal hosted at Glasgow that can *in principal* be used to provide access to sensitive data. This includes three portlets that provide access to Scottish Morbidity Records (SMR) covering hospital admissions; mental health and psychosis data, and death related data. These data sets currently comprise 4-million records and have been released under strict access and usage terms and conditions from the NHS in the UK. A further portlet has been created and offers an accessor function for Census key statistics extracted from MIMAS (Manchester) [19]. A final portlet provides an accessor for geospatial shapefiles hosted by EDINA (Edinburgh) [20]. Figure 2 shows the e-Health portlets in the DAMES portal, which communicate with the Census data server through a GT4 Grid Service. The more general usage and science that is supported by this environment is described in [21], [22].

For the ShinTau evaluation on access to and use of the DAMES portal, a demonstration involving 10 social scientists and software engineers was carried out. The first part of this was to establish an instruction manual describing the process of access and usage of the DAMES portal through ShinTau. We note that given the sensitivity of the NHS data, this scenario focused on the social and

Figure 3. Shintau Linking Service with all three IdPs merged.



geospatial domains, however the solution is generic and can exploit a multitude of attribute authorities. Three Shintau-enabled IdPs were created beforehand, each issuing either an authentication assertion or authorisation attributes according to the following list:

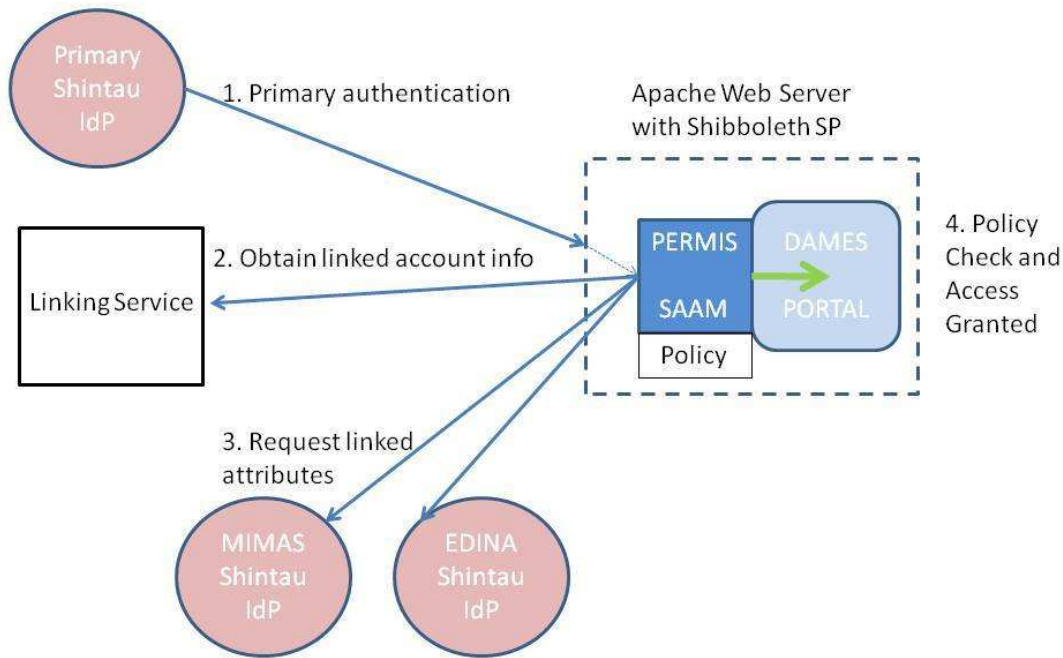
- Glasgow - Primary authentication IdP acting as the home institution;
- Edinburgh - EDINA: Geospatial Data License provider;
- Manchester - MIMAS: Census Data License provider.

At the portal side, the ShinTau-enabled PERMIS SAAM module was deployed and configured with an XML policy which granted access to the portal pages provided the user managed to assert not only an authentication assertion from the Glasgow IdP, but also a geospatial license from EDINA and a Census Data license from MIMAS. Figure 4 shows the flow of user attributes used in the test environment.

The objective of the end-user demonstration was to explore user experiences with configuring and using the Shintau Linking Service so that the conditions demanded by the DAMES portal were satisfied. Each user first logged into the Linking Service using their primary authentication IdP (Glasgow). This allowed them access to the attribute mapping functions of the Linking Service. Each user then added the other two IdPs to their Linking Service mapping account by logging into the EDINA and MIMAS IdPs respectively. Once the three IdPs had been associated, the final step was to choose the SP that this mapping referred to, namely the DAMES portal SP, and commit the mapping to this SP as shown in Figure 3. This ended the user configuration section of the demonstration, and the users then tested the association to see if they could gain access to the secure DAMES portal. Upon visiting the DAMES SP, the user was required to stipulate in a checkbox whether or not they were using ShinTau aggregation on this SP. If this option was not chosen, then when the user logged into the primary Glasgow IdP as normal and was forwarded to the portal, then the PERMIS authorisation policy would establish that insufficient information was provided (i.e. the extra licenses from EDINA and MIMAS were not provided). Repeating



Figure 4. Flow of Information from the Primary IdP which acts as the Authentication Authority, and the Two Extra Attribute Authorities (EDINA and MIMAS) needed to access the Shintau-protected DAMES portal



this part of the demonstration, but having the users select the ShinTau aggregation option, caused the PERMIS module to refer to the two other IdPs which the user had previously linked using the Linking Service and subsequently returns any necessary attributes held about the user. Assuming that the linking of the three IdPs has been done successfully in the Linking Service by the end user, then the DAMES portal security policy is satisfied, i.e. all sufficient licenses are present, and the user may access the portal itself and view the portlet interfaces shown in Figure 2.

As part of this evaluation, a user questionnaire was filled out by most of the participants (9 responded), and the results collated at the University of Kent [23]. The questionnaire queried the user's understanding of the process, its usability, the user's own view on attribute release methods, the software packaging itself and finally some general comments. For each question the user graded each answer on a Likert-type 7-point scale. A summary of the responses revealed the following trends:

- Almost all respondents understood (at least slightly) the principals behind the software;
- All respondents completed the demo without requiring additional assistance;
- All but one of the respondents understood the meaning of each stage in the demonstration;
- Most users (76%) found the Linking Service easy, or

very easy to use;

- Around half felt they could subsequently use the service without the manual;
- A slight majority thought that Shintau should be included with the Shibboleth distribution.

The main themes which became obvious from the responses are that the user-friendliness should be one of the driving forces in any changes to the architecture. The general feeling was that Shintau effectively tackled the problem of federated authorisation. Through the general comments section, many users expressed an interest in being able to choose which attributes from a specific IdP should be asserted and which shouldn't. It should be noted that this kind of operation would break the privacy model of the Linking Service, as at its heart the Linking Service has been designed not to have any knowledge of the contents of attributes sets, it should only know how to reliably assert them.

#### IV. ISSUES

The installation of the ShinTau infrastructure was very complicated, requiring a host of dependent applications to be configured and running before interoperation was possible. To illustrate the dependencies, a ShinTau-enabled IdP requires the following components to be created, configured and started:



- Public Key Infrastructure was set up - Standard issue JANET certificates were obtained;
- Apache Ant, Tomcat (with manager account activated), Axis, Web Server and the latest Java JDK were installed;
- SSL (using the JANET certs) was set up in Apache Web Server, and a ProxyPass directive used to forward requests to Tomcat;
- The Shintau IdP and Liberty Alliance Discovery Service were installed;
- The Shibboleth Security provider was appended to the standard Java security provider list;
- An LDAP server to host the user database was created, with extra schema to handle permisRole and X.509 ACs;
- The ShinTau IdP was configured to use the testshib metadata federation;
- A full Postgres database was installed and configured to store the relations between IdPs and the Linking Service;
- The Liberty Alliance Discovery Service was configured to write/read from this database;
- A full list of all IdP and Linking Service endpoint relations was loaded into the Postgres database.

Because the Shintau infrastructure requires heavily modified/augmented IdPs, the current implementation will not interoperate with the UK Federation (and probably all the other national federations) that are running unmodified Shibboleth version 2. A method for adding a non-ShinTau IdP to the infrastructure was discussed but never tested, involving manipulation of the infrastructure databases.

## V. CONCLUSION

The Shintau software provided by the University of Kent ISSRG group is a standards-based implementation that extends the Shibboleth middleware to allow dynamic aggregation of user attributes and credentials through the creation of a new profile entity known as a Linking Service. This service allows a user to aggregate attributes and credentials from multiple Identity Providers (IdPs), allowing access to services that have access requirements that are not satisfied by the information from a single IdP. This functionality is crucial in the real world where multiple authorities and trust relationships exist.

The application of this infrastructure to a real-life use-case within the ESRC DAMES project was demonstrated at the National e-Science Centre. The DAMES web portal contained portlets that provide services making use of data from external authorities, one a source of census data, and the other a source of geospatial data. In the demonstration, independent IdPs were set up representing these external license providers, plus an institutional IdP. A group of 10 end users successfully used the Shintau software to access a secure portal requiring authentication and authorisation

(licenses) information from three separate IdPs - proving the architecture operates as intended. The users responded to a detailed questionnaire that confirmed that every person who attempted the demonstration was able to link the correct set of attributes in order to access the service, despite layperson knowledge of the technology and its purpose.

Ultimately however, secure access to a portal is only the first stage in realizing the kinds of security models and platforms supported through DAMES. It is highly unlikely that data providers such as the NHS will delegate authorization decisions to a portal hosted at NeSC in Glasgow. Thus whilst they might assert that a user has a license to access a particular data sets, they will also want to make local authorization decisions. The SPAM-GP Attribute Certificate Portlet allows for such scenarios to be realized [24], however, the models of role based access control also need to be augmented with wider security considerations. For example, statistical risk disclosure when dealing with data sets aggregated from multiple providers offers numerous challenges and research possibilities. Many of these are currently being explored within the context of the Scottish Health Informatics Platform ([www.scot-ship.ac.uk](http://www.scot-ship.ac.uk)) project and exploitation of the VANGUARD technology as described in [25].

## ACKNOWLEDGMENT

This work was carried out as a work package under the JISC-funded Shintau Project in collaboration with the University of Kent. The DAMES project is funded by the Economic and Social Sciences Research Council - we gratefully acknowledge their support.

## REFERENCES

- [1] Census.ac.uk <http://www.census.ac.uk>
- [2] Census Area Statistics on the Web, <http://casweb.mimas.ac.uk>
- [3] Data Management through e-Social Sciences, <http://www.dames.org.uk>
- [4] UK Access Management Federation for Education and Research, <http://www.ukfederation.org.uk/>
- [5] D.W. Chadwick and G. Inman, *Attribute aggregation in federated identity management*, IEEE Computer 42(5) 2009, pp. 33–40. (doi:10.1109/MC.2009.143)
- [6] S. Cantor et al., *Shibboleth Architecture: Protocols and Profiles*, Internet2-MACE, <http://shibboleth.internet2.edu/docs/internet2-mace-shibboleth-arch-protocols-latest.pdf>
- [7] The eduPerson Specification <http://www.educause.edu/eduperson>
- [8] D. Chadwick, *Delegation Issuing Service*, NIST 4th Annual PKI Workshop, pp. 62–73

- [9] J. Watt et al., *DyVOSE project: experiences in applying privilege management infrastructures*, In: Cox, S.J. (ed.) Proceedings of the UK e-Science All Hands Meeting 2006, National e-Science Centre, Edinburgh. ISBN 9780955398810
- [10] D.W. Chadwick and A. Otenko, *The PERMIS X.509 Role Based Privilege Management Infrastructure*, Future Generation Computer Systems 19(2) (Elsevier Science BV), 2002, pp. 277-289
- [11] R. Housley et al., *RFC 2459: Internet X.509 Public Key Infrastructure Certificate and CRL Profile*, IETF, 1999, <http://www.ietf.org/rfc/rfc2459>
- [12] D.W. Chadwick, A. Otenko, E. Ball, *Role-Based Access Control with X.509 Attribute Certificates*, IEEE Internet Computing, Mar-Apr 2003, pp. 62-69
- [13] W. Xu et al., *Development of a Flexible PERMIS Authorisation Module for Shibboleth and Apache Server* University of Kent (2005) (doi: 10.1.1.104.3208)
- [14] Shib-Grid Integrated Authorisation (Shintau) Project website, <http://sec.cs.kent.ac.uk/shintau/>
- [15] G. Inman, D. Chadwick and N. Klingenstein, *Authorisation using Attributes from Multiple Authorities: A Study of Requirements*, Presented at HCSIT Summit - ePortfolio International Conference, 16-19 October 2007, Maastricht, Netherlands
- [16] P. Lambert et al., *Utilising a Grid Enabled Occupational Data Environment*, Presented to 16th World Congress of the International Sociological Association (Research Committee 33 on Logic and Methodology in the Social Sciences), Durban S.A., 23-29 July, 2006.
- [17] C. Bayliss et al., *The Design, Development and Application of a Proxy Credential Auditing Infrastructure for Collaborative Research*, submitted to IEEE e-Science 2010, Brisbane, Australia, December 2010.
- [18] K.L.L. Tan et al., *Enabling Quantitative Data Analysis through e-Infrastructure*, Social Science Computer Review 27 (4) 2009, pp. 539-552 (doi: 10.1177/0894439309332647)
- [19] MIMAS, University of Manchester, <http://www.mimas.ac.uk>
- [20] EDINA, University of Edinburgh, <http://www.edina.ac.uk>
- [21] S. McCafferty et al., *e-Infrastructures supporting research into depression, self-harm and suicide*, Phil. Trans. R. Soc. A 368(1925) 2010, pp. 3845-3858. (doi: 10.1098/rsta.2010.0142)
- [22] R.O. Sinnott et al., *E-Infrastructures for Clinical Epidemiological Studies Across Scotland*, Proceedings of e-Society Conference, Algarve, Portugal, April 2008
- [23] G. Inman, D.W. Chadwick, *Summary of the Results of the First User Trial of the Shintau Software Architecture*, <http://sec.cs.kent.ac.uk/shintau/user-trials.pdf> (Accessed 23rd July 2010)
- [24] J. Watt et al., *Tool Support for Security-oriented Virtual Research Collaborations*, Proc. IEEE Int. Symp. on Parallel and Distributed Processing with Applications 2009, pp. 419-424. (doi:10.1109/ISPA.2009.49)
- [25] A. Stell et al., *Designing Privacy for a Scalable Electronic Healthcare Linkage System*, Proc. IEEE Int. Conf. on Computational Science and Engineering 2009, pp. 330-336. (doi:10.1109/CSE.2009.323)